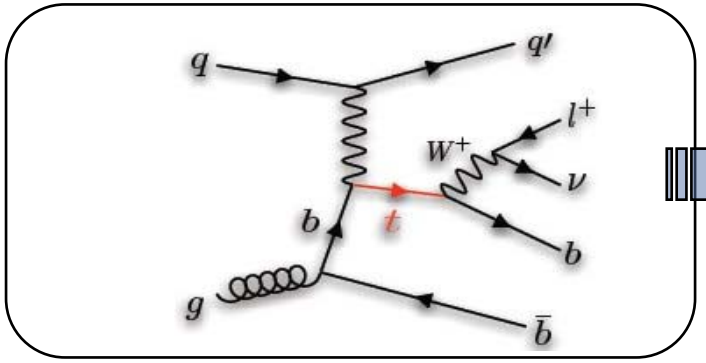
Single-top : from rediscovery to precision measurement

Benoit Clement[1)],Thomas Delemontex[1)], Julien Donini[1)], Cunfeng Feng[2)], Annick Lleres[1)], Arnaud Lucotte[1)], Jin Wang[2)], Xiaohu Sun[2)]

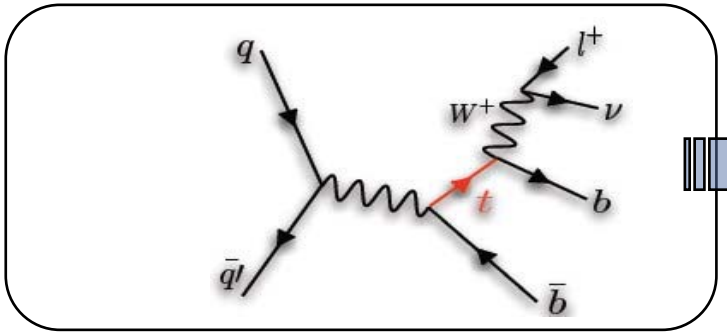
[1)] LPSC Grenoble
[2)] Shandong University

1. Single-top production at hadron colliders & motivations
2. Strategy for single-top cross-section measurement in ATLAS
3. Collaboration Shandong – LPSC in ATLAS

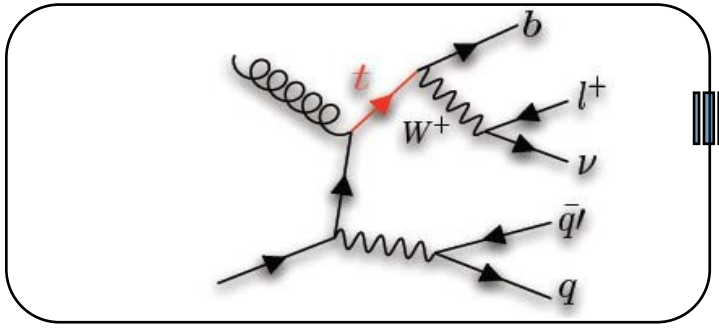
Single-top production @ hadron colliders



– 1 lepton+missing ET (W decay)
– 2-3 high pT jets
– at least 1 b-jet



– 1 lepton+missing ET (W decay)
– 2-3 high pT jets
– exactly 2 b-jets



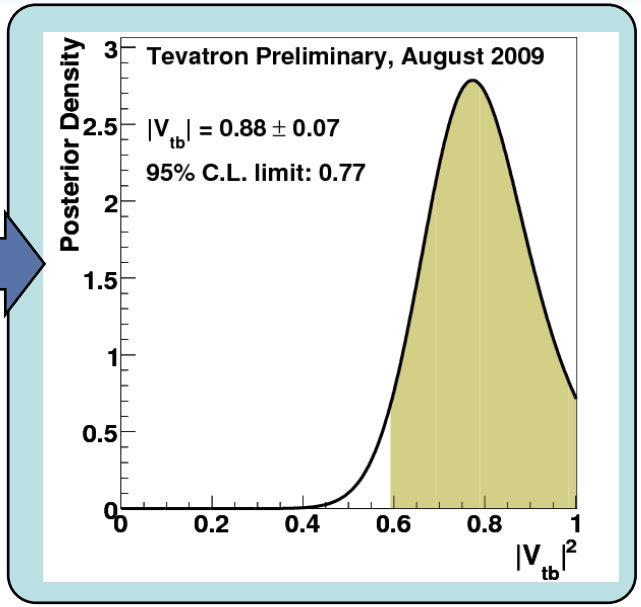
– 1 lepton+missing ET (W decay)
– 2 high pT light jets (W decay)
– exactly 1 b-jet

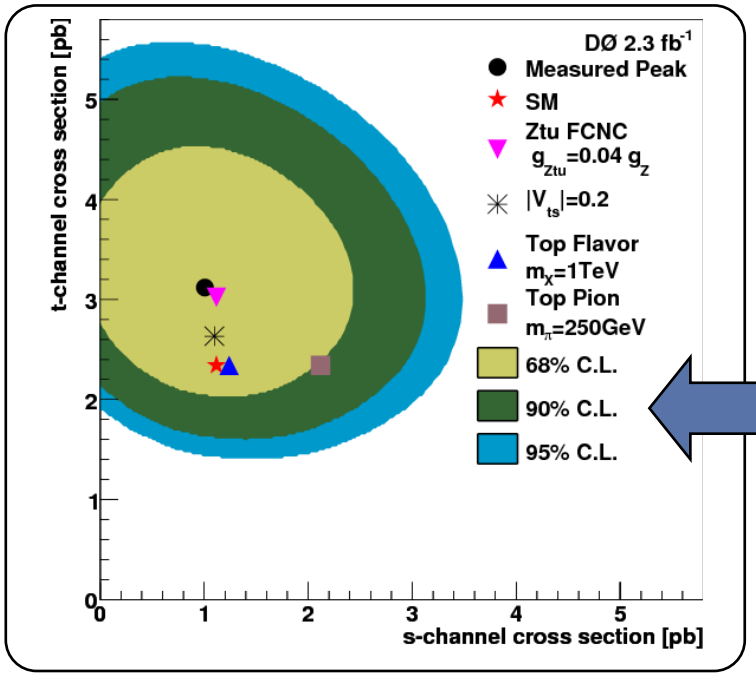
Why measuring single-top processes ?

Constraint on the Standard Model

Precision measurement of the SM

- Establish electroweak production mechanisms of top quark in the t-, s- and W+t channels
- Perform a direct measurement of $|V_{tb}|$, couplings
- Constraint production mechanisms of top quark (single-top events are fully polarized top quarks)





Probe for New physics

Look for deviations in individual cross-sections

Sensitivity to FCNC, anomalous couplings…

→ t-channel events

Sensitivity to extra charged (W+,H+,…) boson

→W+t, s-channel events

Look for new final states :

→ H+t with H+→tb, tau enriched final states

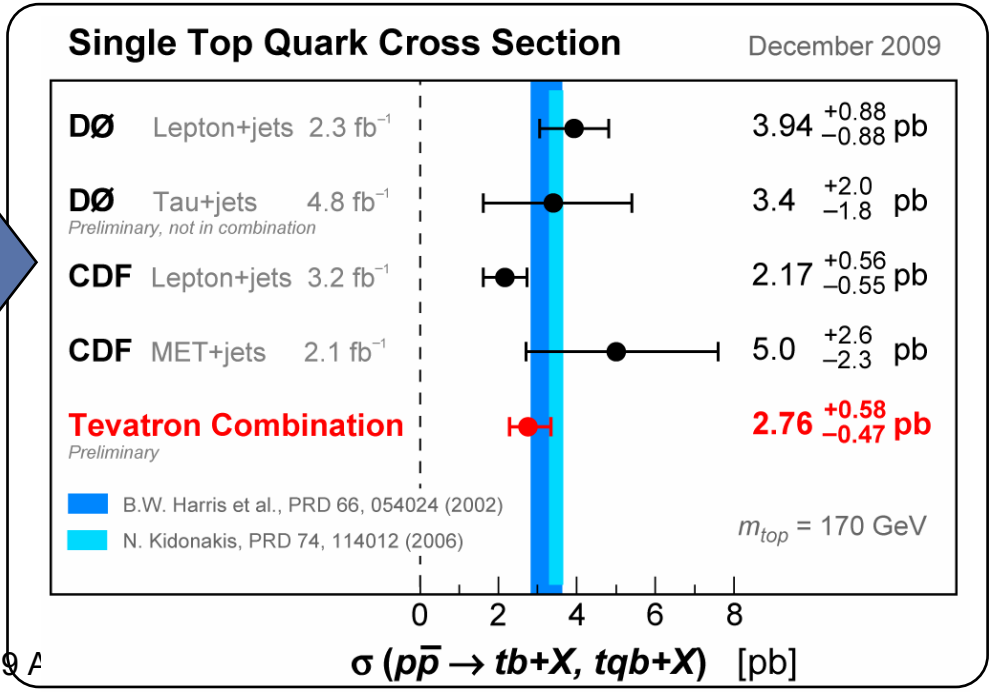
Single-top @ TeVatron

The single-top at the TeVatron

Production established in 2009
- σ(s+t) measured with 21% precision
- $|V_{tb}|$ measured to ~8%

These measurements constitute a ***real challenge***:
- Evidence for t-channel alone still to be shown
- Evidence for s-channel not accessible ?
- W+t is kinematically unaccessible

	DØ 2.3 fb^{-1}	CDF 3.2 fb^{-1}
	Lepton+$\not\!\!E_T$+jets / *b*-tagged	
tb + *tqb* signal *1,*2	223 ± 30	191 ± 28
W+jets	2,647 ± 241	2,204 ± 542
Z+jets, dibosons	340 ± 61	171 ± 15
$t\bar{t}$ pairs *1,*2, *3	1,142 ± 168	686 ± 99
Multijets	300 ± 52	125 ± 50
Total prediction	4,652 ± 352	3,377 ± 505
Data	4,519	3,315



Single-Top @ TeVatron

Why is it so complicated…?

Low initial S/B :

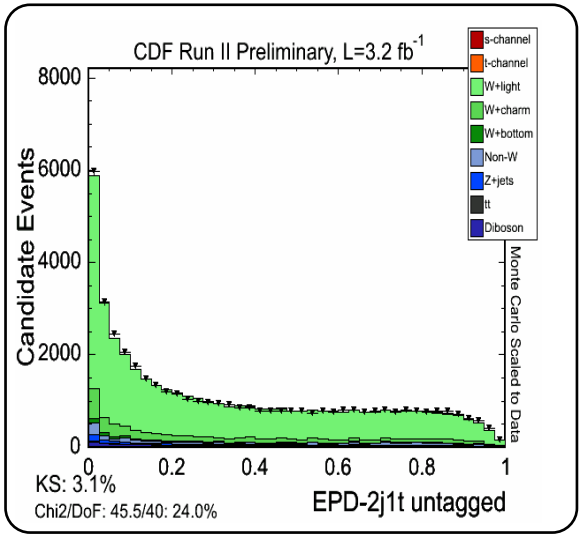
1/200 before b-tagging and ~1/10 after b-tagging

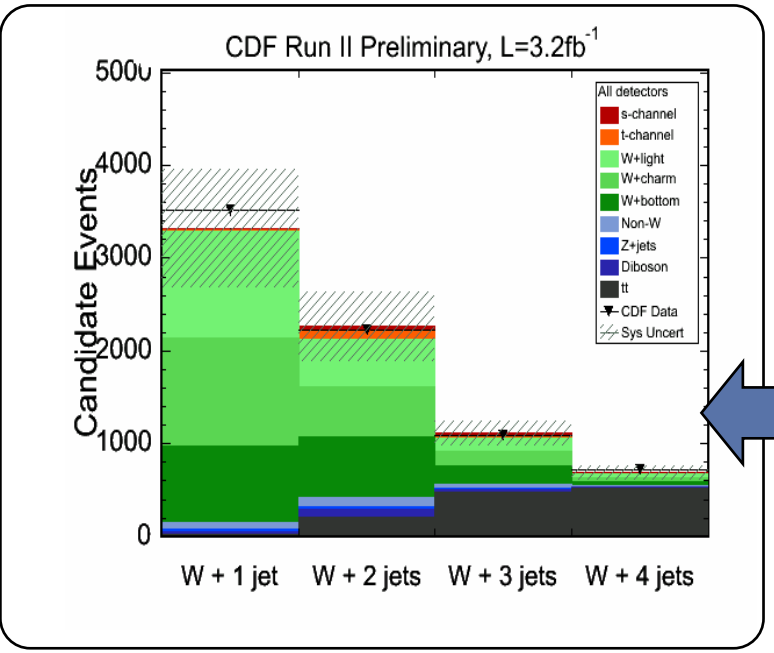
Measurement is complex and requires :

Step 1) a high level of detector understanding

Step 2) a precise determination of background & Modeling

Step 3) the use of sophisticated (MVA) techniques





Step 2) Background Modeling

W+jets determined from data before b-tagging

W+HF processes underestimated by LO generators:

→ K-Factor of 1.4 for Wbb, Wcc, Wcj !

D0 normalize to NLO sample orthogonal to signal

CDF normalizes to W+1jet b-tagged data

ttbar normalized to SM NNLO

QCD from data with non-isolated lepton or low mET

Single-Top @ TeVatron

Why is it so complicated…?

Low initial S/B :

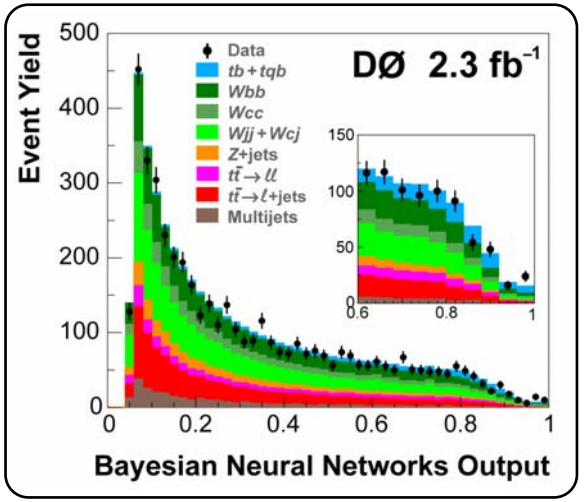
1/200 before b-tagging and ~1/10 after b-tagging

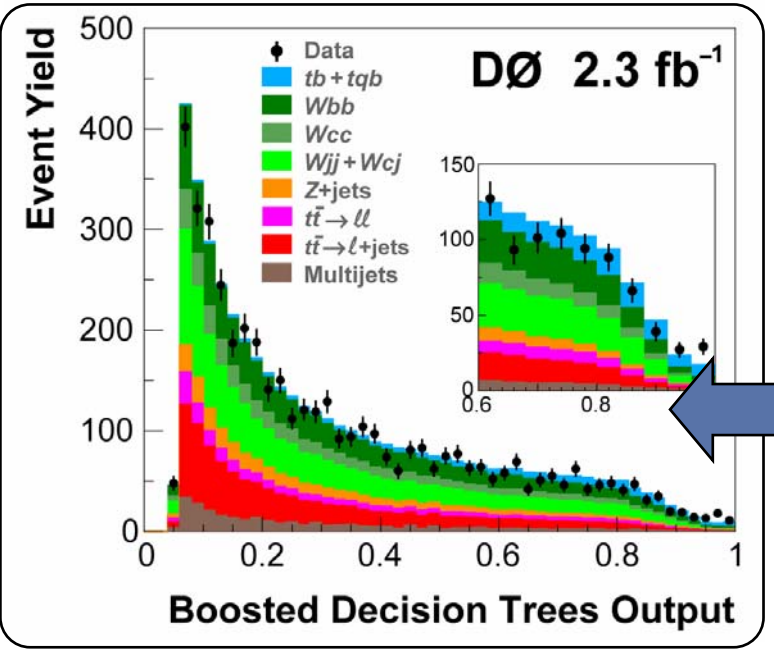
Measurement is complex and requires :

Step 1) a high level of detector bias understanding

Step 2) a precise determination of background & Modeling

Step 3) the use of sophisticated (MVA) techniques





Step 3) MultiVariate techniques

Need to optimize the use of information to discriminate signal from high level of background

- Neural Network
- Boosted Decision Tree
- Matrix Element
- Likelihoods

…and a combination of all those (!)

- Shown to improve significance (60-90% correlation)

Strategy at the LHC…

Single-Top challenge at the LHC

Low S/B, as at the TeVatron

1) Background modeling …now consider:

– W+jets background as next dominant at the pre-selection stage (b-poor backgd)

– Large uncertainties in the W+jets/W+HF jets determination (MC or experimentally)

– top pair background as a dominant background (b-enriched backgd)

→ Uncertainties dominated by the precision with which background is known

2) Need to use data driven techniques to understand the backgrounds

– to constraint W+jets and W+HF jets production

– to determine top-pair cross-section precisely in signal phase space

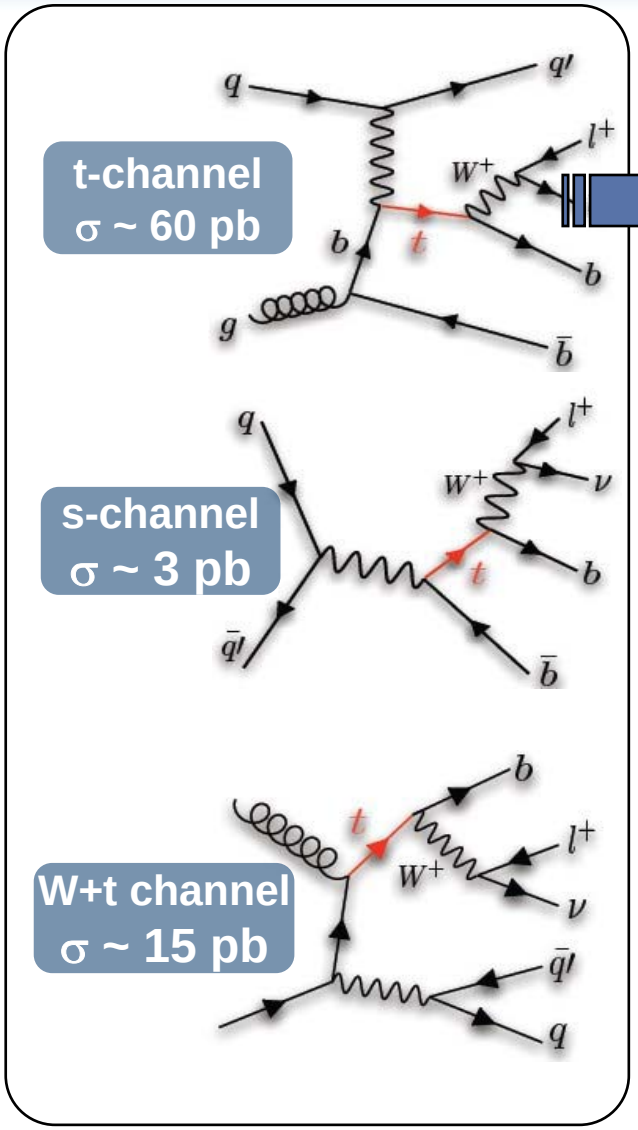
→ Validation of the techniques used is a major issue

3) Need to use sophisticated techniques to isolate signal from background

– use of MultiVariate analysis techniques

→ Validation of these techniques is then a major issue

Strategy in ATLAS @ 7 TeV



Common pre-selection

Inclusive lepton trigger ~80% efficiency
- at least one isolated high p_T lepton
- at least two jets
- at least one b-tagged jet
- missing ET

Background modeling

- Background dominated : S/B ~few % after b-tagging
- Systematics on background dominate
  - → Use of data driven techniques mandatory
  - → Necessity to enhance purity : use of MVA

Optimization of signal selection

MVA vs CutBased selections
→ First validate MC vs Data on whole phase space
Selection optimization:
- Cuts on MVA outputs that minimize systematics
- Use of toy MC to generate D,B as Poisson and D,B, ε for all sources of systematics

Cross-section extraction using $\sigma = D-B/\varepsilon L$

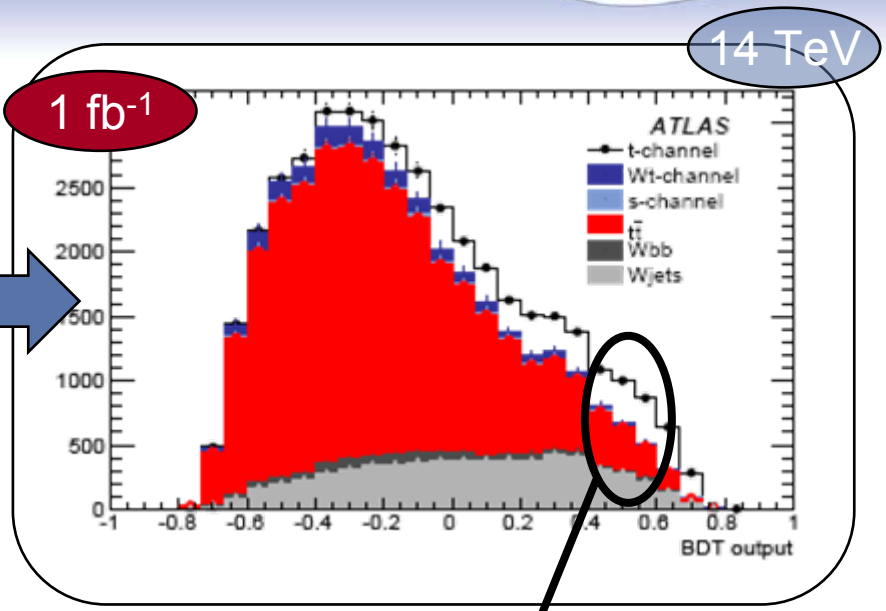
Single-top at the LHC : t-channel

Event Selection

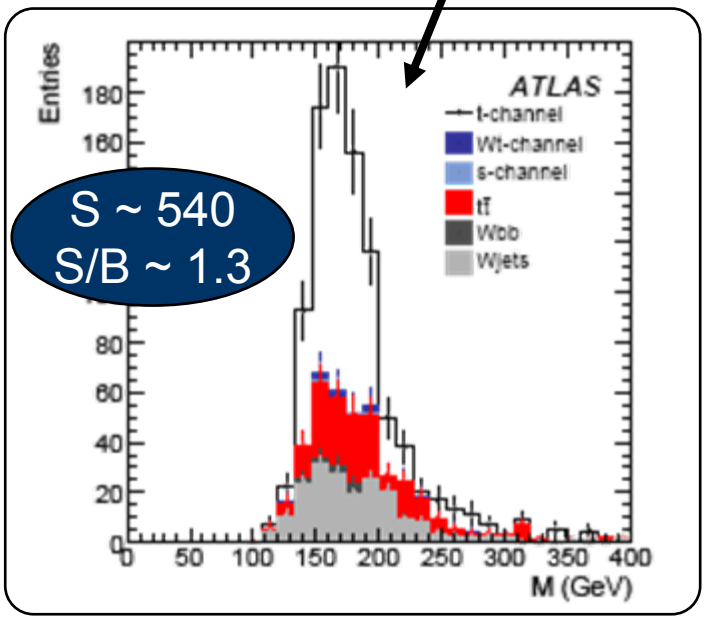
Exactly 2 high-p_T jets:
- 1 high pT central b-jet
- 1 forward light jet $|\eta|>2.5$

BDT analysis
- 6 variables not (too) sensitive to JES
- Optimization vs top pair

Signal : $\varepsilon \approx 1\text{-}2\%$ N ~ 500 evts



Systematics @ 1 fb^{-1}	δσ/σ
Luminosity	8.8%
Jet energy scale	9.9%
B-tagging	6.6%
Backgrounds (MC)	8.2%
ISR/FSR + PDF …	9.9%
MC statistics	7.9%
Total SYSTEMATIC	**22.4%**
Total STATISTICAL	**5.7%**



Single-top at the LHC : W+t channel

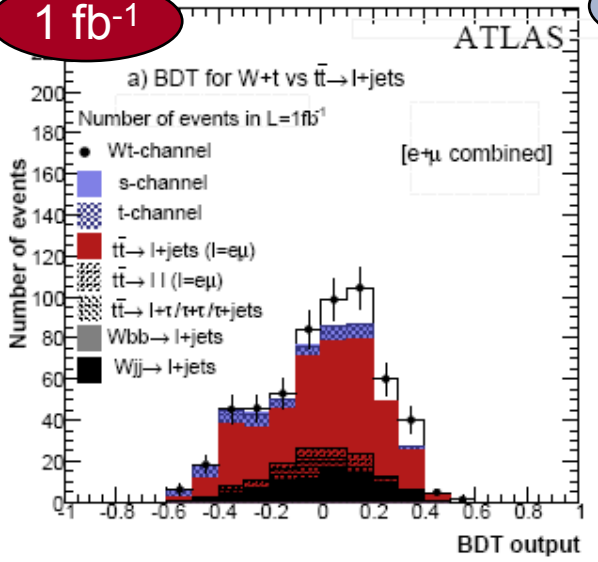
W+t channel selection

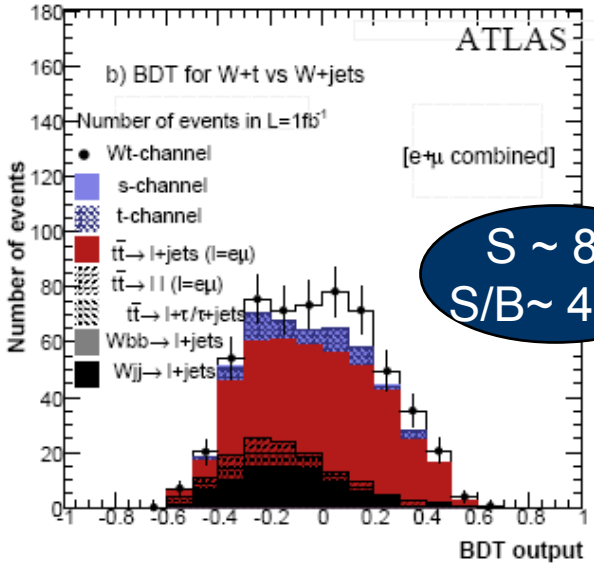
2 high-p_T untagged-jets
reconstruct W→ jj
Veto of a 2nd b-tagged jet
Boosted Decision Trees:
– Set of 4 BDTs vs specific bkgds
– Set of BDTs for 2/3/4 jet final states

Systematics @ 10 fb^{-1}	δσ/σ
Luminosity	7.9%
B-tagging	6.6%
Jet energy scale	2.0%
Backgrounds (MC)	9.6%
ISR/FSR+ PDF +b frag	13.3%
Lepton ID, trigger	6%
Total SYSTEMATIC	19.4%
Total STATISTICAL	6.6%

1 fb^{-1}

14 TeV



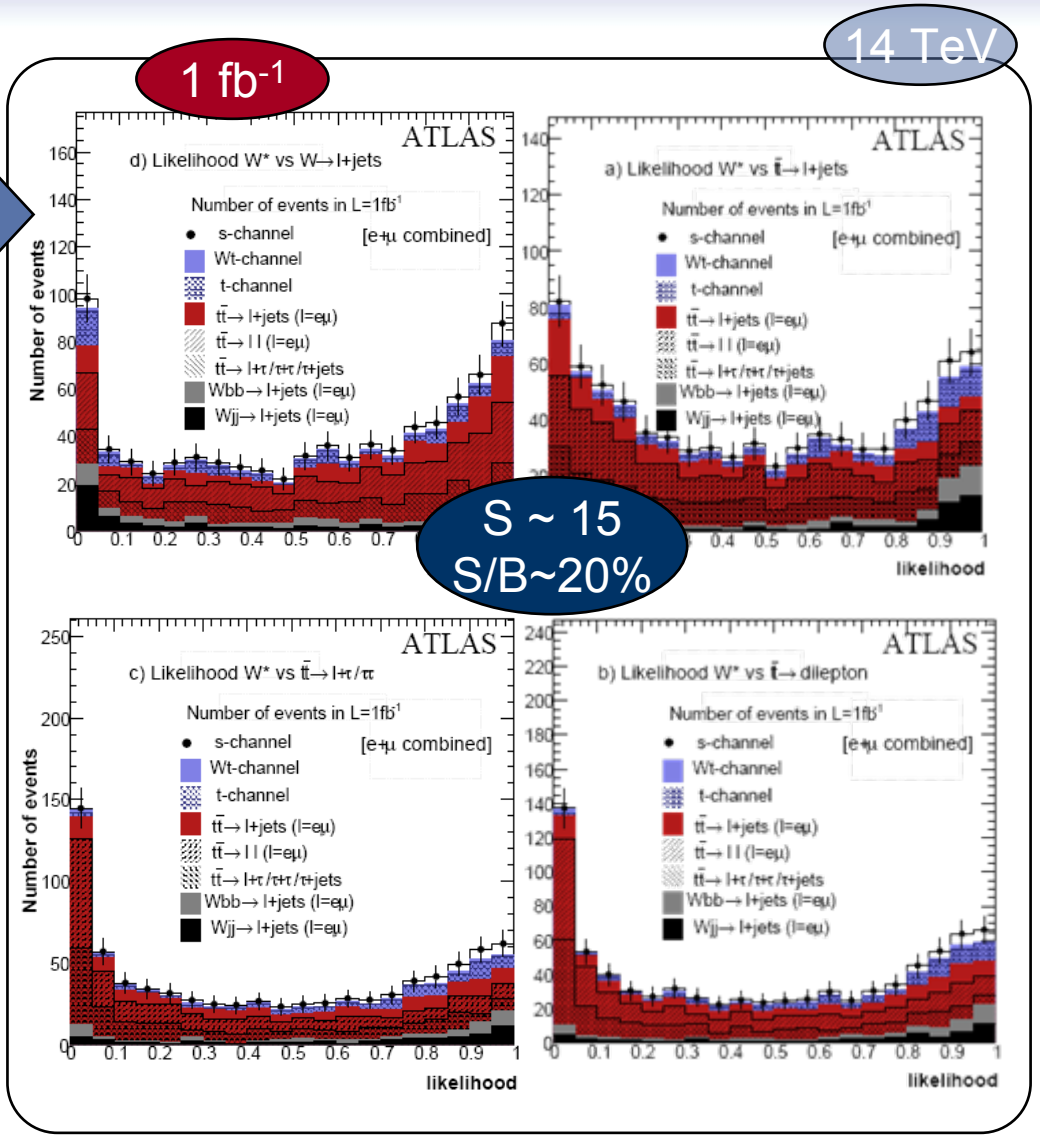


S ~ 80
S/B~ 40%

Single-top at the LHC : s-channel

s-channel selection
– 2 high-pT b-jets
– Veto of a 3rd jet
Likelihood functions:
– 5 likelihoods vs specific bkgds
– Choice of ~indpt variables

Systematics @ 10 fb^{-1}	δσ/σ
Luminosity	18%
B-tagging	25%
Jet energy scale	8%
Backgrounds (MC)	16%
ISR/FSR+ PDF +b frag	30%
Lepton ID, trigger	6%
Total SYSTEMATIC	48%
Total STATISTICAL	20%

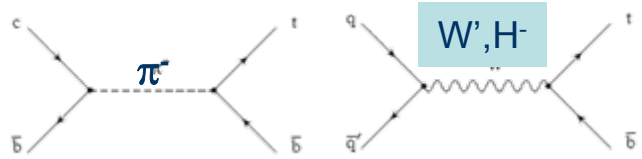


Single top@ LHC as probe to New Physics

Interpretations Beyond SM

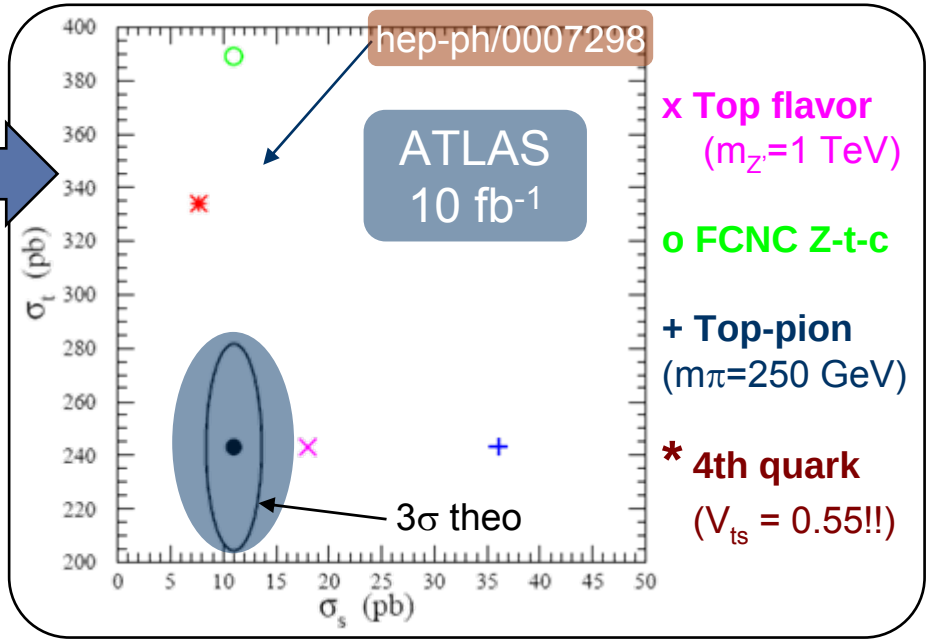
Single tops as probe to NP
- Cover a large range of NP in top production or decay
- via couplings or new particles

s-channel sensitive to:
- W' in GUT/ED
- $H^{\pm}$ in NMSM or MSSM
- Techni-pion



t-channel sensitive to:
- Anomalous couplings
- Anomalous polarization

W+t channel sensitive to:
- $H^{\pm}$ search !
  pp → $H^{\pm}t$ production

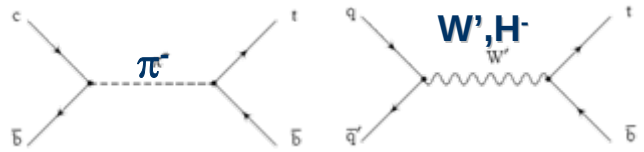


Single top @ LHC as probe to New Physics

Interpretations Beyond SM

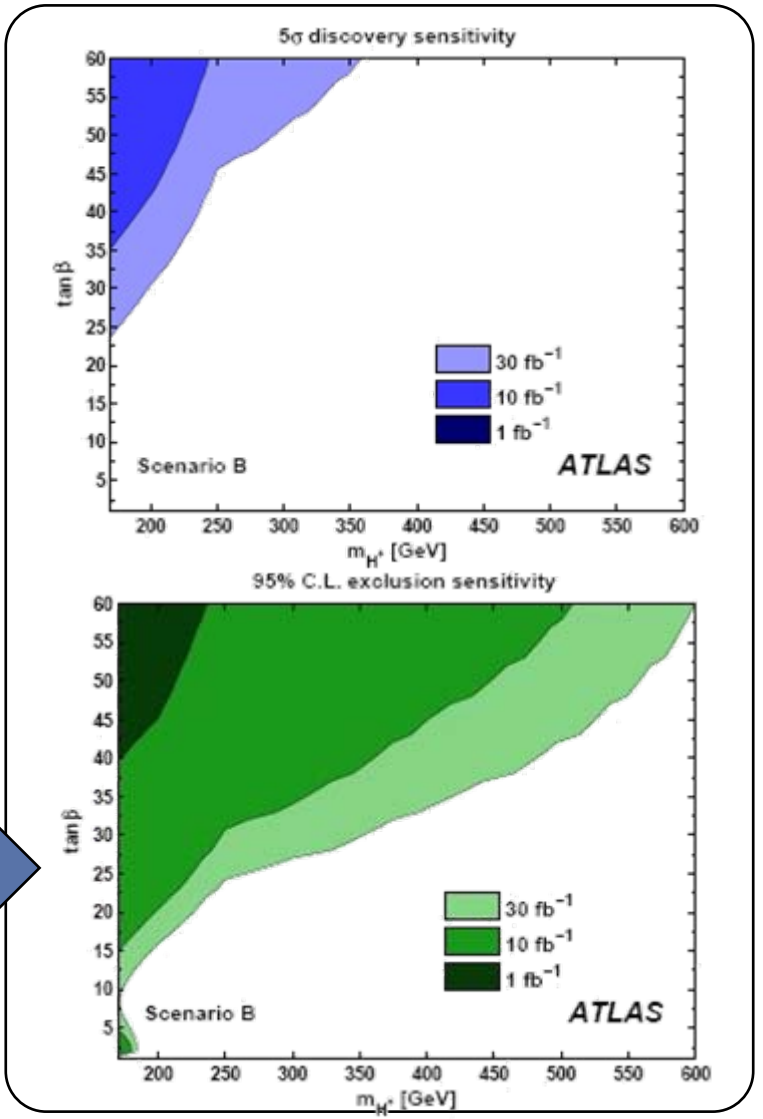
Single tops as probe to NP
- Cover a large range of NP in top production or decay
- via couplings or new particles

s-channel sensitive to:
- W' in GUT/ED
- $H^{\pm}$ in NMSM or MSSM
- Techni-pion



t-channel sensitive to:
- Anomalous couplings, FCNC
- Deviation in polarization

W+t channel sensitive to:
- $H^{\pm}$ search !
  pp → $H^{\pm}t$ production



Collaboration Shandong-LPSC

Three main axes of the collaboration

1) Objet calibration : B-tagging for single-top selection
   - Calibration of the b-tagging efficiency
   - Calibration of the b-tagging rejection rate

2) Preparation for single-top analyses
   - Estimate of the W+jets background from data
   - Measurement of the t-channel analysis
   - Search for the Wt-channel

3) Signal Modeling with Monte Carlo generator:
   - Validation/Production of NLO single-top channels

The development of common tool/expertise

1) b-tagging with a common infrastructure, fitting tools etc…
2) Validation of MC involving both groups
3) Common analysis framework initally developed at the LPSC

The modality of the collaboration

1) One PhD student sharing its time (50%/50%)in Shandong & Grenoble
2) Presence of a Master student in Grenoble. Maybe a future PhD ?

Collaboration Shandong-LPSC

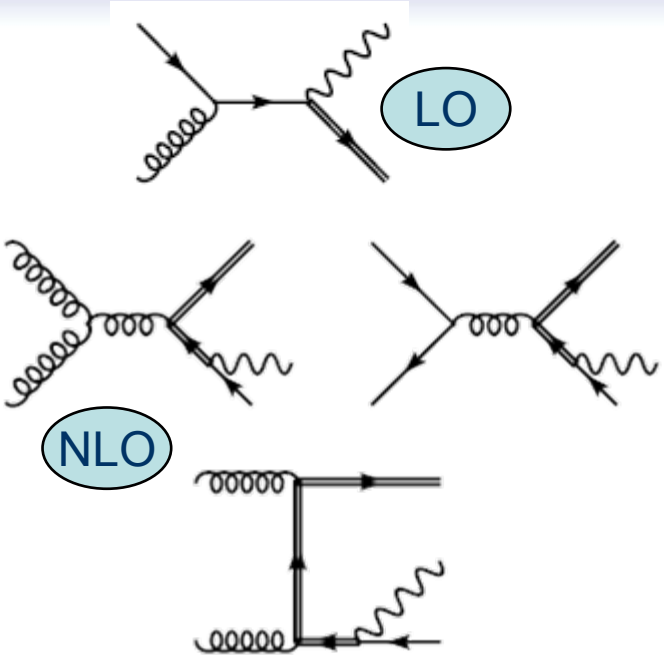
1) Monte Carlo Validation

NLO generators are available for Single-top signals

→ Use of MC@NLO and POWHEG generators

Understand the issues with generators:

- Tuning of Hadronization/showering, pile-up
- Dependence in PDF set
- Validation of NLO generators



Validation of the W+t channel as an example

Single-top channel distinction possible at LO

→NLO corrections make things more tricky, e.g. interference with ttbar diagrams

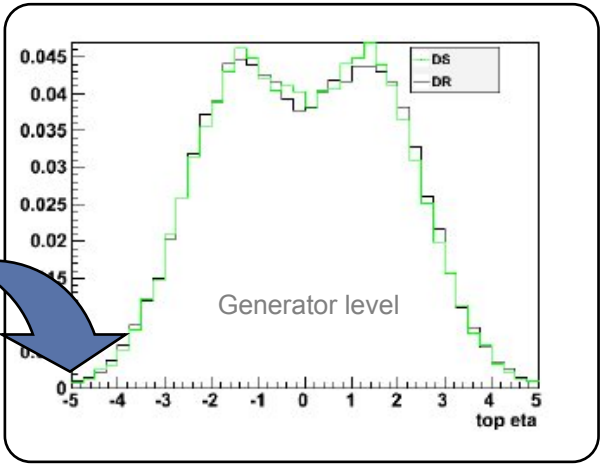
Procedures to differenciate tt and Wt *theoretically*

- Diagram Subtraction scheme
- Diagram Removal scheme

arXiv 0908.0631

Experimental ways to define these 2 *experimentally*

- Cut on second b-jet pT, etc…
- →Check consistency of the experimental approach with the two schemes



Collaboration Shandong-LPSC

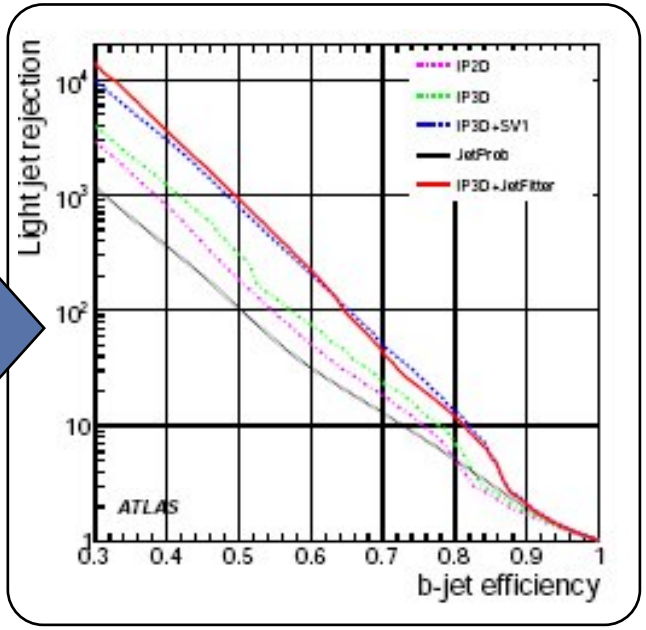
2) B-tagging

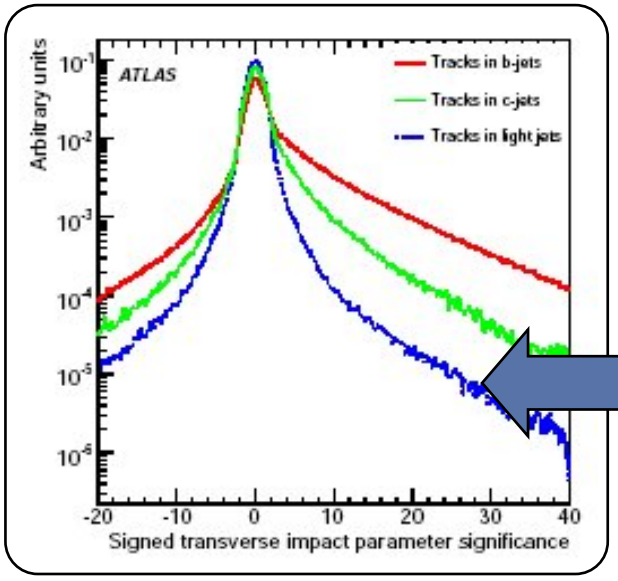
B-tagging is a mandatory element

- reject W+jets, QCD etc…
- Select 1-2 b-tagged jet final states

Calibration of b-tagging using data thus crucial:

- Use of pTrel measurement
- Use of System8 techniques





Mistag rate determination

High rejection rate are needed

Method to calibrate the rejection rates using data

- Use negative part of secondary vertex taggers from QCD-enriched data
- →Backport the results to pre-selected sample (W+jets)
- →using a Transfert Matrix

Summary and conclusion (1)

Road Map for Single-Top Physics

May be a long way to reach comparable sensitivity to TeVatron

1) Rediscovery of top quark
   - Benchmark for detector understanding
   - →leptonID, jet, jet energy scale, missing transverse energy, b-tagging, (c-tagging ?), mistag rate…

2) Understanding of main background processes
   - QCD and W+jets determination
   - W+HF jets with devoted measurements
   - Top pair in signal acceptance

3) Cross-section measurements
   - Need *well understood* data and modeling
   - Optimization of the available information : MVA
   - Validation of inputs to MVA techniques
   - Combination of different techniques

4) Reach at 7 TeV for 1 fb^{-1}
   - First LHC run *may allow* single-top rediscovery in the t-channel
     - will be both systematics and stat. limited
   - Sensitivity to W+t channel @ 7 TeV is not yet established
     - Relevant since a channel not accessible @ TeVatron
   - s-channel is clearly a longer term measurement

Summary and conclusion (2)

Collaboration Shandong-LPSC

Organization aimed at sharing expertise

- Common tasks, common tools/ infrastructure developed
- 2 PhDs working together in both Grenoble & Shendong + 5 seniors
- → 1 PhD + 1 master student from Shandong so far

Directed towards most relevant features with early data:

- Calibration of B-tagging performance
- Understanding of W+jets/QCD backgrounds
- NLO generators validation

Group focused on the main single-top production

- Search for single-top t-channel
- Sensitivity for the W+t

Backup

Strategy in ATLAS with the early data…

Single-top t-channel with early data

Dominant production of single-top

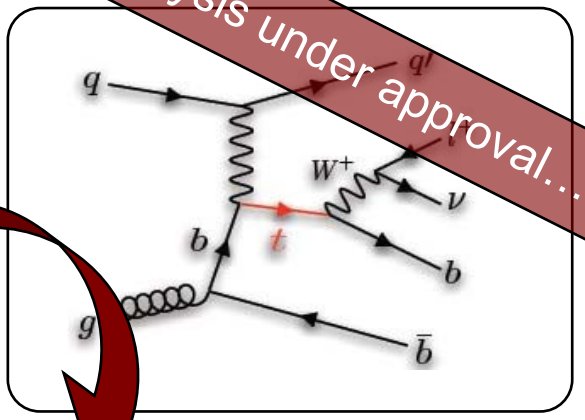
Data driven techniques to estimate background

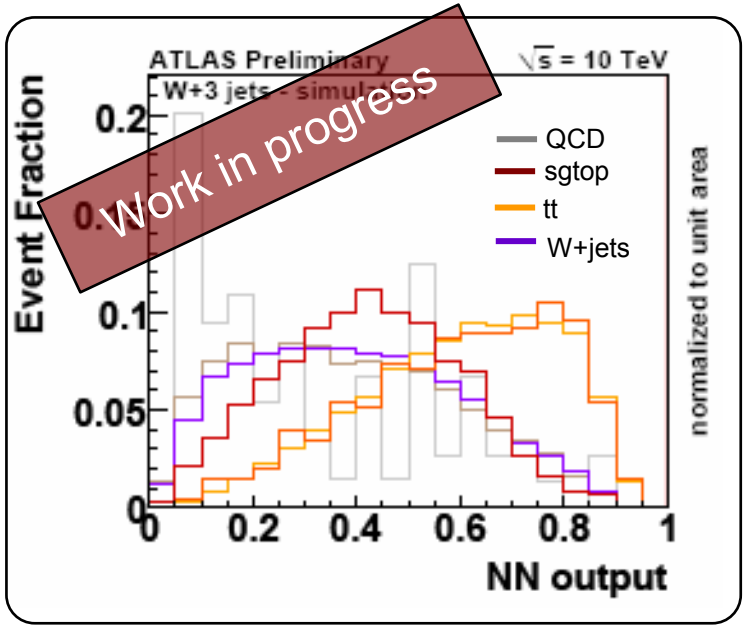
- Determine QCD with matrix method & monitor
- Determine backgrounds from data

Use robust Likelihood with well understood variables

- Optimize selection vs systematics

Analysis under approval…





Work in progress

Background modeling

- Use preselected 3-jets events, before b-tagging
- Construct a NN discriminant, fit W+jets and ttbar rates
  → orthogonal sample, kinematically similar to signal region
- Take into account systematics associated to shape and efficiencies

→ Expected precision (stat+syst) using ensemble tests

Evidence should be reached w/ ~300-400 pb^{-1} at 10 TeV

Strategy in ATLAS with the early data…

Single-top t-channel with early data

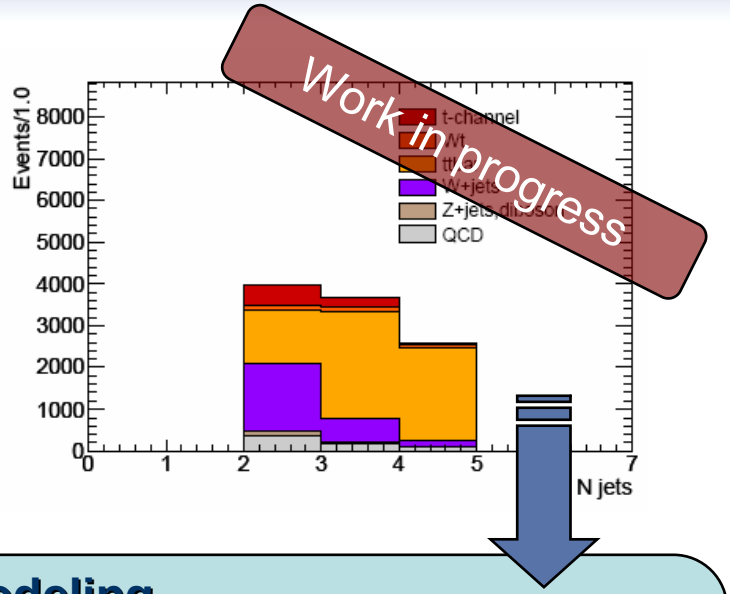
Dominant production of single-top

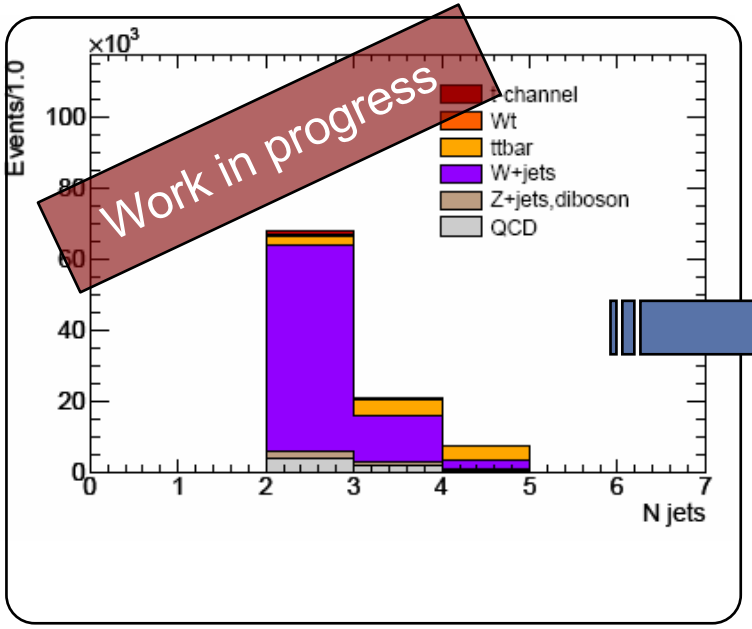
Data driven techniques to estimate background

- Determine QCD with matrix method & monitor
- Determine backgrounds from data

Use robust Likelihood with well understood variables

- Optimize selection vs systematics





Background modeling

- Use preselected 3-jets events, before b-tagging
- Construct a NN discriminant, fit W+jets and ttbar rates
  → orthogonal sample, kinematically similar to signal region
- Take into account systematics associated to shape and efficiencies

→ Expected precision (stat+syst) using ensemble tests

Evidence should be reached w/ ~300-400 pb^{-1} at 10 TeV

Strategy in ATLAS with the early data…

Single-top t-channel with early data

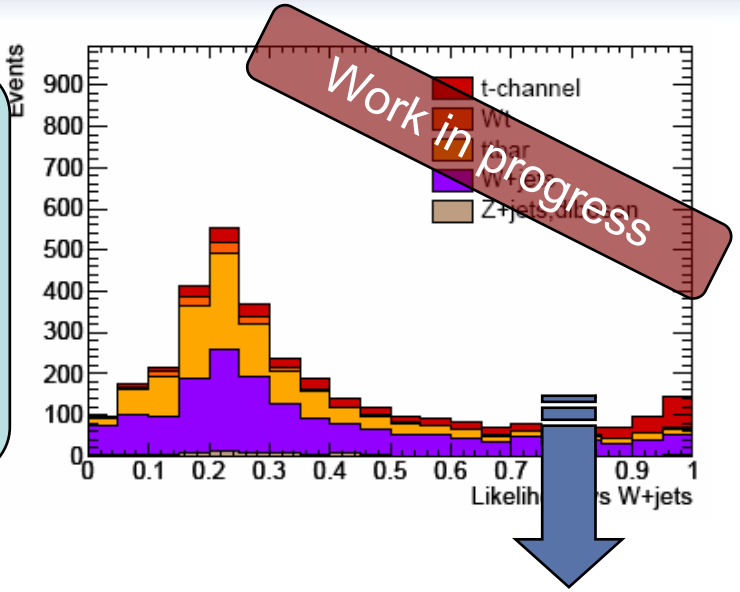
Dominant production of single-top

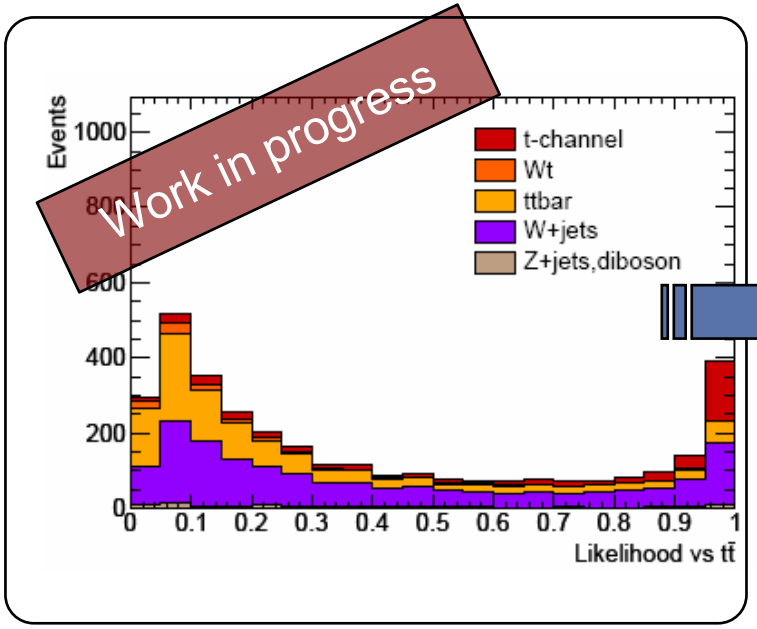
Data driven techniques to estimate background

- Determine QCD with matrix method & monitor
- Determine backgrounds from data

Use robust Likelihood with well understood variables

- Optimize selection vs systematics





Build a likelihood to enhance purity

- Use preselected 2-jets btagged events
- Use a few angular/robust variables

Take into account systematics associated to shape and efficiencies

→ Expected precision (stat+syst) using ensemble tests